②

D50

COPY NO. ____

ELECTROCORTICAL ACTIVITY AND OPERATOR WORKLOAD:

A COMPARISON OF CHANGES IN THE ELECTROENCEPHALOGRAM AND IN EVENT-RELATED POTENTIALS

JUNE 1981 MDC E2427

FINAL REPORT

PREPARED BY:

KIRMACH NATANI
FRANK E. GOMER

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

MCDONNELL DOUGLAS ASTRONAUTICS COMPANY · ST. LOUIS DIVISION
Box 516, Saint Louis, Missouri 63166 (314) 232 0232

MCDONNELL DOUGLAS CORPORATION

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont)

LIST OF PAGES

LIST OF FIGURES

PREFACE

The research described herein was supported by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency (Contract MDA-903-78-C-0181) and by Internal Research and Development funds from the McDonnell Douglas Astronautics Company-St. Louis Division. Dr. Craig Fields, Director of the Cybernetics Technology Office, sponsored the work for the Government. Dr. Frank E. Gomer and Dr. Kirmach Natani of the McDonnell Douglas Astronautics Company-St. Louis Division were Program Manager and Principal Investigator, respectively.

1.0 ABSTRACT

Pilot workload considerations are beginning to have a distinct cognitive emphasis, due to changes that are occurring in display formats and operating procedures. Consequently, new techniques for assessing workload must be developed that are more sensitive to fluctuations in attentiveness and in the capacity to time-share among several system demands for processing, decision-making, and action.

A flight simulation experiment was conducted in which pilots were trained to follow a commanded flight profile and maintain airspeed while concurrently performing threat avoidance and target acquisition tasks. Two types of brain electrical activity, event-related potentials (ERPs) and the "ongoing" electroencephalogram (EEG), were analyzed in response to direct manipulations of difficulty in executing commanded flight maneuvers. Because of these manipulations, the difficulty of time-sharing was influenced indirectly. Both workload level and continued experience with the flight simulation tasks affected the magnitude and latency of the P_{300} component of the ERP, as well as the distribution of EEG power. These findings support the concept of multidimensional processing resources and the distinction between controlled versus automatic modes of processing.

The results of this investigation are encouraging. They suggest that a physiological method used to evaluate mental workload in laboratory settings may be applied successfully, in at least some instances, to situations in which individuals perform operationally relevant tasks.

2.0 INTRODUCTION

Adaptive systems requiring real-time interactions between an operator and a computer must depend on manual or voice inputs. The sophistication of these interactions would be enhanced if electrophysiological signals which reflect the current status[1] of the operator served as input to the computer. In a previous report, Gomer, Beideman, and Levine (1979) reviewed the benefits of increasing air crew effectiveness through the development of such a "biocybernetic" communication channel. They outlined several options that could be executed by a computer if, following analysis of electrophysiological signals, it were deemed necessary to unburden or assist the pilot.

Before biocybernetic techniques can be implemented in an operational environment, however, it must be shown that changes in electrophysiological activity, e.g., event-related brain potentials (ERPs), are reliably associated with changes in mental functions. Donchin and his colleagues (Isreal, Wickens, Chesney, and Donchin, 1980b) have reported recently that the size of the P_{300} component of the auditory ERP is a sensitive measure of residual processing capacity during visual task loading. These investigators simulated an air traffic control task of monitoring symbolic representations of aircraft which moved in straight-line paths across a display. The observers were required to respond to sudden changes in the displayed trajectories of certain aircraft. Mental workload was manipulated by varying the number of aircraft to be monitored simultaneously. Auditory probes were presented in the background as a Bernoulli series of low- and high-pitched tones. The high tones occurred quite infrequently, and the observers were instructed to maintain a running count of these events and to report the count at the end of the experiment. It was found that the P_{300} component of the high-tone ERP exhibited a monotonic reduction in area as the number of displayed aircraft was increased from zero to eight. While these and other laboratory investigations (cf. Thatcher and John, 1977) have provided promising demonstrations of the relationship between changes in brain electrical activity and changes in mental

[1] We define changes in "operator status" as momentary fluctuations in attentiveness and in the capacity to time-share among multiple demands for information processing and decision-making.

function, the demands imposed upon the subjects have been somewhat constrained from an operational point-of-view.

Gomer (in press) concluded that electrophysiological studies must be transitioned at this time from the laboratory to appropriate part-task and full-mission simulation. Due to problems with artifacts in the electrophysiological data and the complexity of the signal analysis involved, he also recommended that biocybernetic applications be restricted to those air crew tasks which are critical to the success of the mission or occur during period of heavy workload. Under these conditions, analyis of electrophysiological activity may provide a means of anticipating imminent deteriorations in pilot proficiency.

2.1 Statement of the Problem - ERPs are transient voltage fluctuations generated in the brain during discrete events, such as during the registration and processing of sensory information or during decision-making. These transient voltage fluctuations are embedded in the more prominent ongoing electrical activity of the brain, making it necessary to use signal averaging techniques for their extraction. The spatial and temporal characteristics of ERP waveforms serve as profile information that is essential to the evaluation of mental function. Additional profile data are available in the form of power spectra of the continuous electrical activity of the brain. This continuous activity is termed electroencephalographic (EEG) activity, and changes in spectral intensity within restricted frequency bands are presumed to be indicative of changes in complex mental functions (see Gevins and Schaffer (1980) for a review of the literature).

2.2 Purpose and Scope - The work described here addresses some of the significant problems which exist in the area of electrophysiological data collection and analysis in operational settings. We have conducted a feasibility study, employing a part-task flight simulation, to determine if it is possible to obtain ERP and EEG data which may be useful for communication purposes. Real-time biocybernetic applications were not attempted in this study. Rather, the primary purpose was to manipulate pilot workload during a reasonably demanding simulated mission and obtain electrophysiological data for off-line reduction and evaluation. Future studies must perform real-time analysis of electrophysiological data for closed-loop applications.

As noted above, most ERP (and EEG) studies have been conducted in university laboratories. In this setting, subject posture and movement, as well as task demands, are highly controlled to minimize the presence of artifacts. Unfortunately, a recent, more operationally-oriented investigation of the relationship between workload level and ERPs during flight simulation tasks focused upon early components of the waveform that are sensitive to stimulus properties rather than to cognitive processes (Spyker, Stackhouse, Khalafalla, and McLane, 1971). Therefore, our intent was to replicate, during flight simulation, the finding by Isreal et al. (1980b) of an inverse relationship between the processing requirements of a primary task and the size of the P_{300} component of the auditory ERP elicited by discrete probes presented in the background.

2.3 Part-Task Simulation - Concurrent pilot tasks were used, including: (1) following a commanded flight profile (through adjustments in pitch attitude and bank angle) while maintaining a constant airspeed; (2) monitoring for auditory threat warnings, with switch activation required for threat avoidance; and (3) tracking, acquiring, and designating a symbolic target on a CRT display. The simulation assumed an attack aircraft flying an air-to-ground strike. Horizontal and vertical Attitude Director Indicator (ADI) bars simulated pitch and bank angle commands that an aircraft's mission computer would provide to steer the pilot to the target area. Target positions and speed across the CRT were chosen to simulate a ground-stabilized sensor system.

3.0 METHODS

3.1 Subjects - Five subjects participated in this study. All were McDonnell Douglas Corporation (MDC) employees. They were 19 to 45 years of age, with a mean age of 32.6 years. Prior experience for the group ranged from no previous flight or simulator time to over 100 hours solo flight time, in one case, and over 500 hours of simulator time, in another case. Two individuals had previous experience with electrophysiological studies. One of the five subjects was left handed.

3.2 Experimental Apparatus - The apparatus consisted of a fixed-base, part-task simulator configured in the forward cockpit of an RF-4B mockup (Figure 3-1). A PDP-12 computer (Digital Equipment Corporation) was programmed to provide the control laws and system dynamics. This computer also collected, processed, and stored pilot performance data.

9 2601

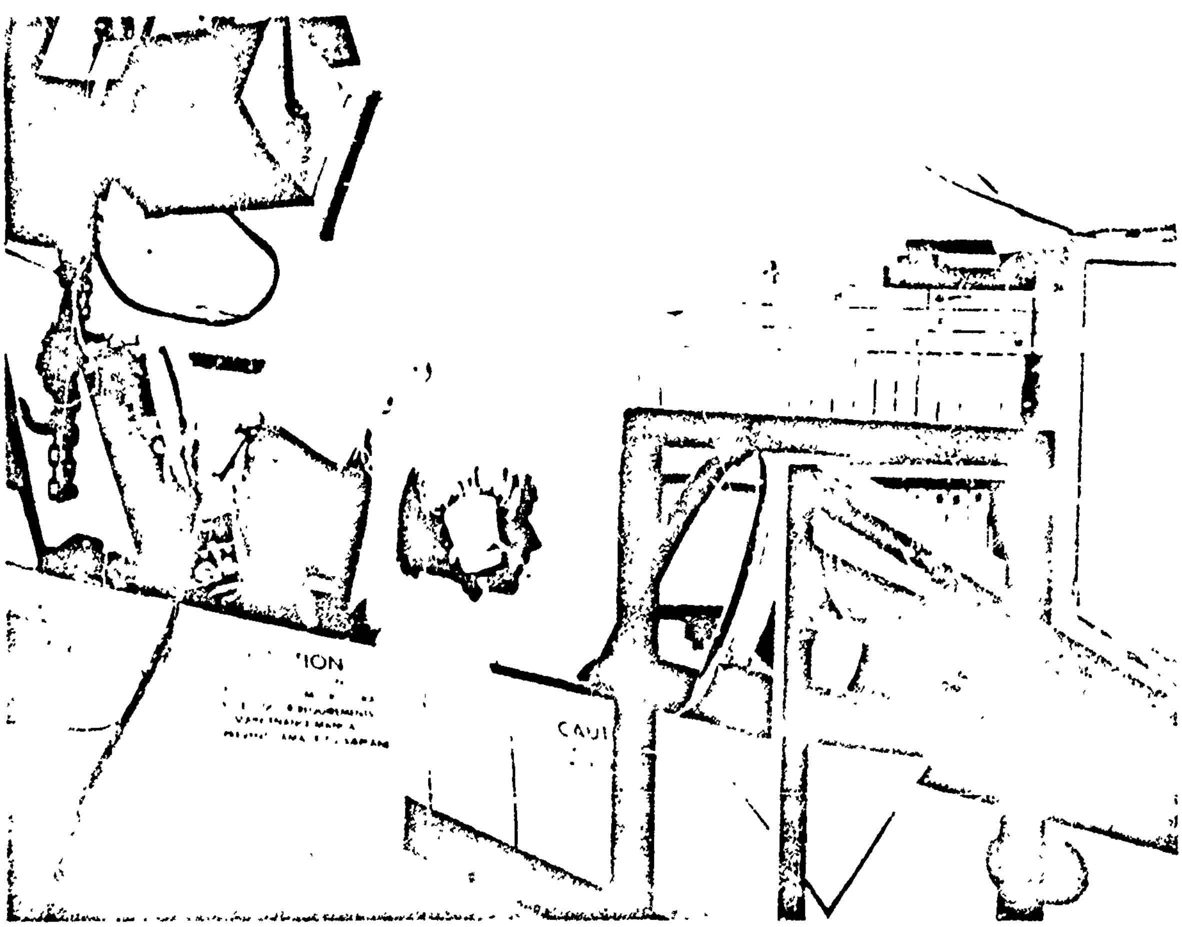

FIGURE 3–1 FLIGHT SIMULATOR USED IN THIS INVESTIGATION

The active displays of the simulated crew station consisted of: (1) a Lear Siegler, Inc. Attitude Director Indicator (ADI); (2) a panel-mounted meter labeled to provide airspeed information in knots/hour with an indicated range of 100 to 800 knots; and (3) a CRT display (Phillips Telequipment Oscilloscope, Model S54A) to provide simulated target acquisition information in the form of a fixed cross hair and moving target symbology. Nonfunctional (for this simulation) instruments enhanced the physical characteristics of the crew station. The approximate distance from the subject's eye position to the instrument panel was 28-30 inches.

The functional controls included a flight stick assembly (right hand), a throttle (left hand), and a finger-operated device for two-dimensional target tracking. The later device was a force transducer that was integrated into the throttle quadrant. A thumb-operated threat response switch was also a part of the throttle assembly. Finally, a switch for designating target position was mounted on the flight stick assembly.

3.2.1 <u>Control/Display Functions</u> - Flight control and target tracking output voltages were analog-to-digital (A/D) converted to allow comparisons with commanded pitch attitude, bank angle, and airspeed, as well as with desired target symbol position. Absolute deviations were calculated to assess pilot performance and to provide analog feedback voltages to the ADI, airspeed, and target displays. All flight and target tracking dynamics were comprised of first-order components. The flight dynamics equations (representative of those used in the A-7) were written with the following assumptions: (1) the aircraft was trimmed (pitch); (2) the aircraft was coordinated (yaw); (3) drag was a simple linear function of airspeed; (4) bank angle held when the stick was released; and (5) all control was accomplished with pitch, roll and throttle adjustments.

The input to the ADI had two features: (a) the actual pitch and bank angle of the aircraft, and (b) the difference between the actual attitude and the commanded attitude to be flown by the pilot. The first input was displayed through the relationship of a fixed aircraft symbol and the attitude sphere of the ADI. The second input appeared in the form of lateral and vertical displacements of the bank and pitch steering bars, respectively, relative to the aircraft symbol.

Input to the airspeed indicator was a function of both throttle setting and pitch attitude. Thus, a forward stick position (pitch down) increased airspeed and an aft stick position (pitch up) decreased airspeed with no change in throttle position.

The computer directed the dynamics of the target symbol and generated the fixed cross hair on the CRT display. The target symbol showed the target in proper perspective to the aircraft (i.e., with regard to airspeed, pitch and bank angle) on a horizontal plane. The acquisition cross hair was fixed at the center of the CRT.

A two-axis isometric controller integrated into the throttle provided left-hand control over the target symbol when it was presented on the CRT display. The target symbol could be moved to any location on the CRT screen with this fingertip force transducer. Target "acquisition" required movement of the target symbol over the fixed cross hair and then depression of the switch mounted on the flight stick.

The computer also controlled the presentation of discrete tonal events which represented threat warnings. The pulses which coded the tone presentations provided triggers to initiate ensemble averaging (off-line) of electrophysiological data. When a threat warning was perceived by the pilot, he activated a switch on the throttle to simulate the selection of countermeasures.

Analog voltage outputs from the flight control stick, throttle and the target tracking device were sampled by the computer twenty times per second, i.e., at 50 msec intervals. The input signals to the displays were also generated by the computer at an update rate of twenty times per second.

3.3 Electrophysiological Data Collection - Three channels of EEG and one channel of EOG data were amplified with a Beckman Accutrace 200 clinical EEG system. All signals were recorded with 1 second time constants and were low-pass filtered at 30 Hz (18 dB/octave) to reduce movement and 60 Hz artifacts.

The EEG electrodes were positioned on the head according to the 10-20 system. Recordings were made from three sites: 1) the midpoint between C_z and P_z (coded $\overline{C_zP_z}$); 2) P_3 on the left side; and 3) the homologous site, P_4, on the right

side. Linked ear lobe leads were used as the reference for the active sites. Beckman silver EEG electrodes were used with a Beckman bentonite paste. Electrode impedances were not allowed to exceed 5,000 ohms and in most instances were below 1,000 ohms. Beckman Biopotential electrodes, filled with Beckman cloride paste, were used to obtain the EOG. The EOG electrodes were placed laterally on the outer canthus and above the right eye to record both horizontal and vertical components of eye movements. These signals were used to reject individual ERPs containing large eye movement artifacts during the ensemble averaging process. A conservative criterion was adopted for acceptable EOG variance within the recording epoch.

EEG data from which ERPs were extracted were recorded on FM magnetic tape. Off-line digitization, editing, averaging, and analysis were performed by a PDP 11/40 computer. The sampling interval was 4 msec. Resultant ensemble average ERPs were 800 msec in duration, with a 100 msec pre-stimulus baseline.

Separate analyses were performed on the EEG data. The EEG recorded across an entire simulation trial (150 sec) was high-pass filtered digitally to reject all frequencies below 2.34 Hz. The EEG was then Fourier-analyzed to compute estimates of spectral intensity in θ (3.5 - 7.5 Hz) and α (8.0 - 12.0 Hz) frequency bands. A resolution of 0.39 Hz was adopted. Spectral intensities also were determined for resting/eyes closed and resting/eyes opened conditions just prior to testing.

3.4 Procedure - As noted above, the primary task required the pilot to follow a commanded flight profile through adjustments in pitch attitude and bank angle and to maintain a constant airspeed. A secondary task consisted of short-term compensatory tracking (slewing) to "acquire" the computer-driven target symbol displayed on the CRT. An ancillary auditory vigilance task was also imposed in the form of a threat monitoring task. Trial duration was 150 seconds and intertrial interval, marked by a red light on the crew station instrument panel, was 3 minutes. The experimental trial time line is illustrated in Figure 3-2.

3.4.1 Attitude Control - The subjects had to continuously follow pitch and bank angle commands displayed on the ADI. The command inputs required compensatory tracking, with the provision that a displayed deviation in one direction be corrected by movement of the flight control stick in that direction. Attitude

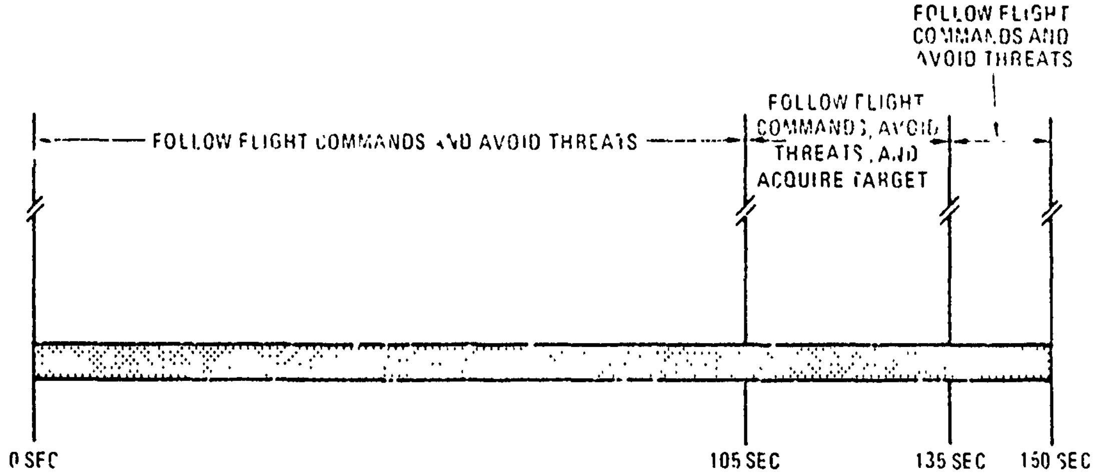

FIGURE 3–2 EXPERIMENTAL TRIAL TIME LINE

error was measured as root-mean-squared (RMS) in both axes. RMS error, in this case, is defined as the standard deviation of the periodically sampled absolute differences between actual and commanded attitude.

3.4.2 Airspeed Control - Subjects were required to maintain airspeed by throttle manipulation. During testing, RMS airspeed errors were recorded. These error values were based on the difference between the displayed airspeed and a commanded airspeed of 300 knots.

3.4.3 Secondary Target Tracking - Finger-controlled, two-dimensional tracking of a target was required of the subject in this task. The principal measure of target tracking performance was acquisition time. Acquisition time was defined as time in seconds from target appearance on the display until the subject centered the target symbol over the fixed cross hair and simultaneously depressed the acquisition button on the flight control stick. The subject could not predict the initial position of the target nor the trajectory it would follow. Each target was displayed for a maximum of 30 seconds and moved at a constant speed (on the display) of 0.25 in./sec.

3.4.4 Threat Avoidance - Throughout each experimental trial, a total of thirty high and low frequency tones was presented to the pilot to simulate auditory threat warnings. The high tone (1600 Hz) was presented frequently (probability of occurrence equal to 0.67) and represented radar "painting" activity.

Detection of a high tone did not require a response. On the other hand, the low tones (450 Hz) represented a "missile launch" threat that did require activation of a switch mounted on the throttle to simulate the selection of countermeasures. The low tones served as "rare" events in Donchin's (1979) probe stimulus technique for generating auditory ERPs. All tones were embedded in a background of white noise. Both tone stimuli produced ERPs, but, as noted above, only the "rare" events required a response.

3.4.5 Manipulation of Workload - Workload was equated with the difficulty of following pitch and bank angle commands while maintaining airspeed. To manipulate difficulty, pitch and roll disturbances were combined with pitch and roll forcing functions. The same pitch and roll disturbances (diagrammed in Figure 3-3), which

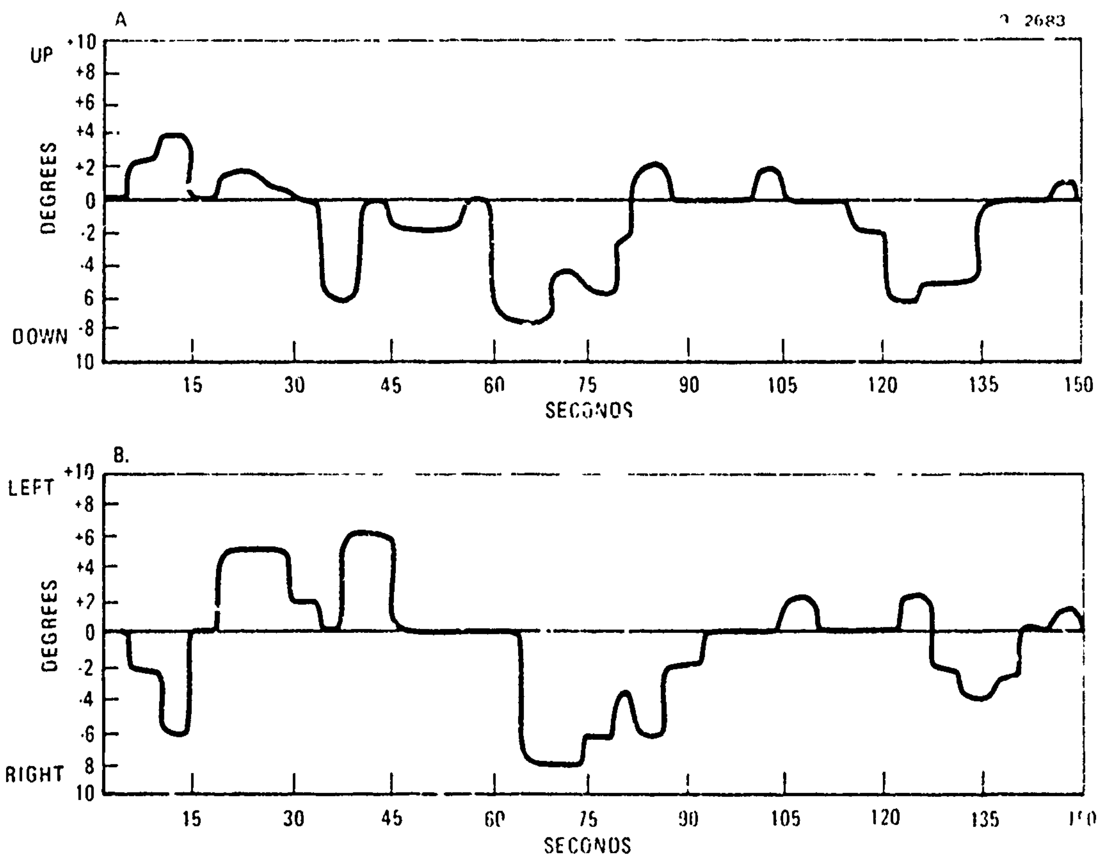

FIGURE 3-3 PITCH (A) AND ROLL (B) DISTURBANCES USED DURING THE MISSION TO REPRESENT NOMINAL GUST CONDITIONS

represented nominal gust conditions, were used for low and high workload conditions. The low and high workload pitch forcing functions are diagrammed in Figure 3-4, whereas the low and high workload roll forcing functions (two each) are diagrammed in Figures 3-5 and 3-6, respectively.

Subjective estimates of workload are frequently determined for complex operational tasks (Johannsen, Moray, Pew, Rasmussen, Sanders, and Wickens, 1979). Therefore, the two workload levels selected for this study were based upon prior ratings by experienced military pilots (Warner, Drennen, and Curtin, 1976).

3.4.6 Design - We wanted to evaluate the effects of the different treatments on performance of the primary task, i.e., on the accuracy of pilot adjustments in attitude and airspeed. We also wanted to analyze acquisition times for secondary target tracking and validate our manipulation of workload level, since it has been reported that increases in primary task difficulty produce increases in the time taken to perform a secondary task (Isreal et al., 1980b; Ogden, Levine, and Eisner, 1979). Therefore, the experimental design was a repeated measures factorial design representing: (a) workload level (high vs. low); (b) test session (first or second day); (c) successive blocks of experimental trials within a session (trials 1 to 10, 11 to 20, or 21 to 30; and (d) subjects (N=5). For the EEG recordings, electrode placement ($\overline{C_zP_z}$, P_3, or P_4) was added as a variable in the analysis. For the ERP data, threat avoidance stimulus (high vs. low tone) was also included as a variable. A Statistical Analysis System (SAS) computer program was used for the analyses of variance and Newman-Keuls tests performed on the data.

Each subject received three hours of practice on the simulated mission prior to testing. This practice was distributed across two different days. Participants were briefed concerning the purpose of the study. The requirements for following a commanded flight profile and maintaining airspeed, for secondary target tracking, and for monitoring of threat warnings were described in detail and demonstrated before practice trials were initiated. During practice, the subjects were encouraged to seek clarification on any aspect of the study. Electrophysiological data were not recorded during practice trials.

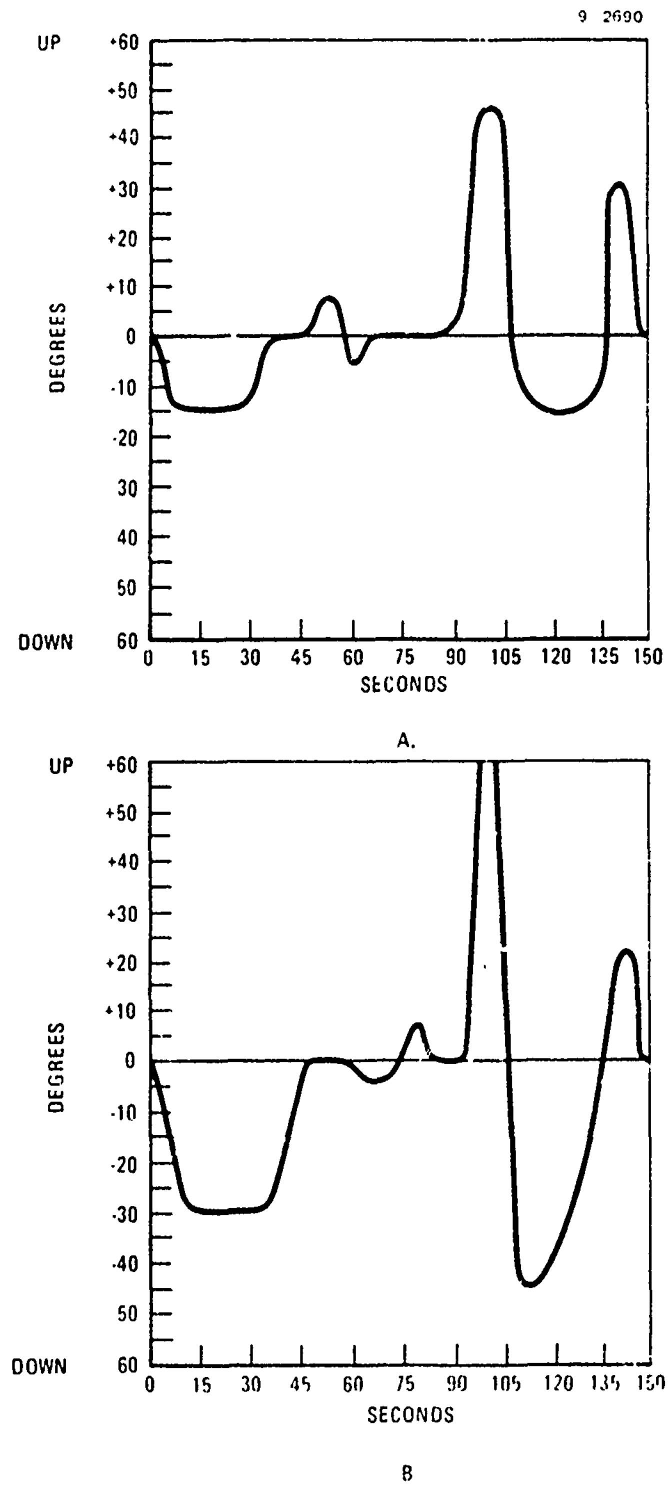

FIGURE 3–4 LOW (A) AND HIGH (B) WORKLOAD PITCH FORCING FUNCTIONS

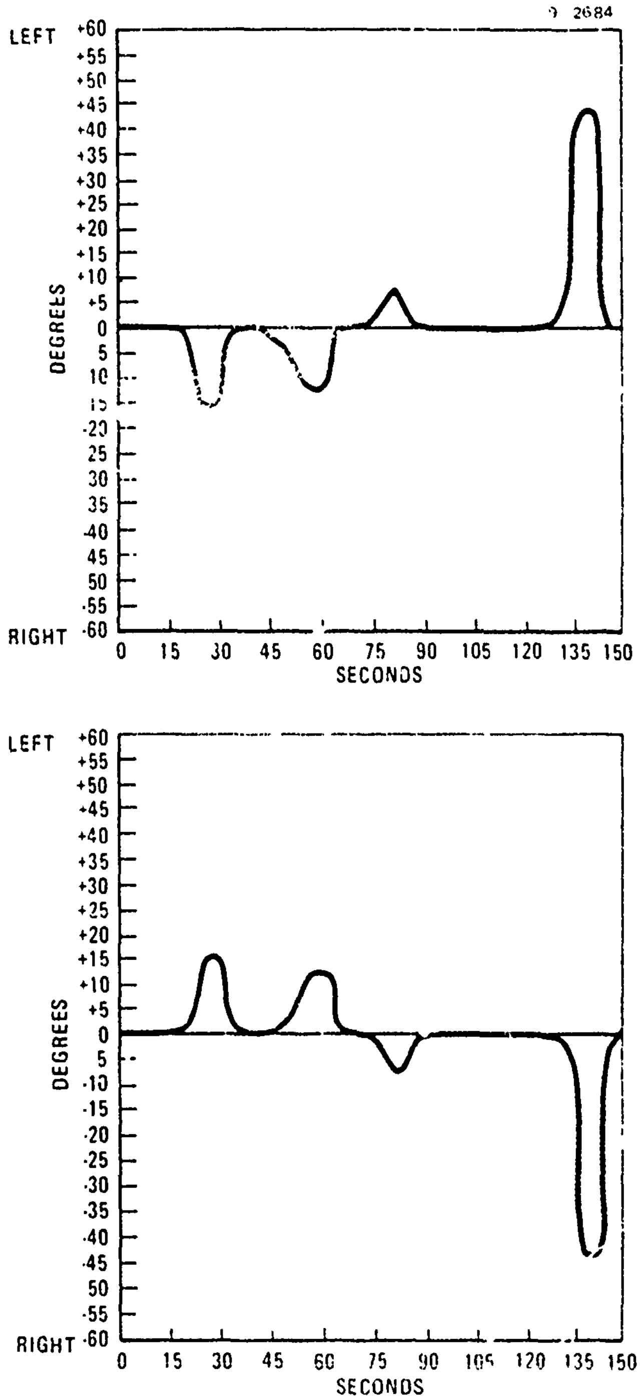

FIGURE 3–5 TWO LOW WORKLOAD ROLL FORCING FUNCTIONS PRESENTED RANDOMLY ACROSS TRIALS WITHIN A SESSION

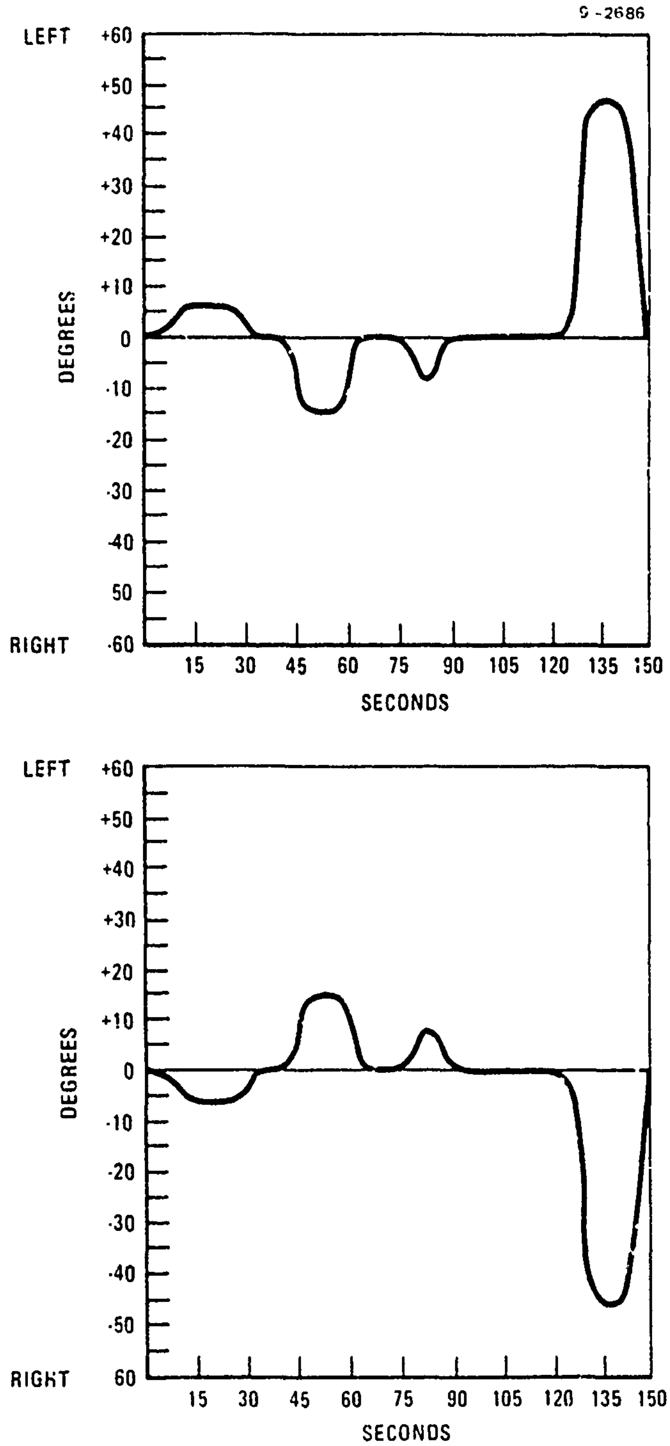

FIGURE 3–6 TWO HIGH WORKLOAD ROLL FORCING FUNCTIONS PRESENTED RANDOMLY ACROSS TRIALS WITHIN A SESSION

Following the completion of practice, the subjects received two days of testing. Each test day consisted of 30 trials, 150 seconds in duration. Within each block of 10 trials, high and low workload trials were presented randomly. This prevented the subjects from developing stereotypical response patterns. The time line shown in Figure 3-2 indicates that for the first 105 seconds, the subject was responsible for flight control and threat monitoring. Before the trial began, the aircraft "was frozen in space." At time zero, the subject assumed control of the aircraft and followed steering commands to the simulated target area. Target tracking and acquisition took place between 105 and 135 seconds, followed by egress from the simulated target area. Threat monitoring was superimposed during the latter two phases of the mission as well.

4.0 RESULTS AND DISCUSSION

4.1 Accuracy in Following a Commanded Flight Profile and Maintaining Airspeed - RMS errors were computed separately for pitch attitude, bank angle, and airspeed. However, a composite measure of accuracy was derived to reflect the overall difficulty of following a commanded flight profile and maintaining airspeed (cf. Williges and Wierwille, 1979). Several weighting functions were developed to combine the three dependent variables. Because of the high degree of similarity in the results of the analyses performed on the different composite scores, we present our findings for a simple summation of the three types of error data. This approach has been effective in the past for evaluating the difficulty of executing reasonably simple instrument flight maneuvers (Knoop and Welde, 1973).

Mean composite errors were determined for each block of ten trials within a session and entered into the analysis of variance. These data, pooled across subjects and trial block within a session (first ten vs. second ten vs. third ten trials), are shown in Figure 4-1. The main effects of test session ($F = 6.34$, $df = 1/44$, $p < 0.02$) and workload level ($F = 4.12$, $df = 1/44$, $p < 0.05$) were significant. That is, there was a substantial reduction in composite RMS error from day 1 to day 2. In addition, composite RMS error was greater for the high workload than for the low workload condition. However, planned comparisons revealed that while composite RMS error was clearly greater for the high workload than for the low workload condition during test session 1 ($p < 0.01$), no significant difference was apparent in this measure for the two workload levels during test session 2. By the second test session, the pilots had become more proficient in following a commanded flight profile and in maintaining airspeed, and, on the basis of the error data, it would appear that the demands associated with the high workload and low workload conditions were now quite comparable. It is conceivable, however, that the pilots were merely expending greater effort to perform the primary task as accurately on high workload trials as they did on "less demanding" low workload trials.

4.2 Target Acquisition Times - To further evaluate our manipulation of task demands, we assumed that if greater resources were indeed required of our pilots

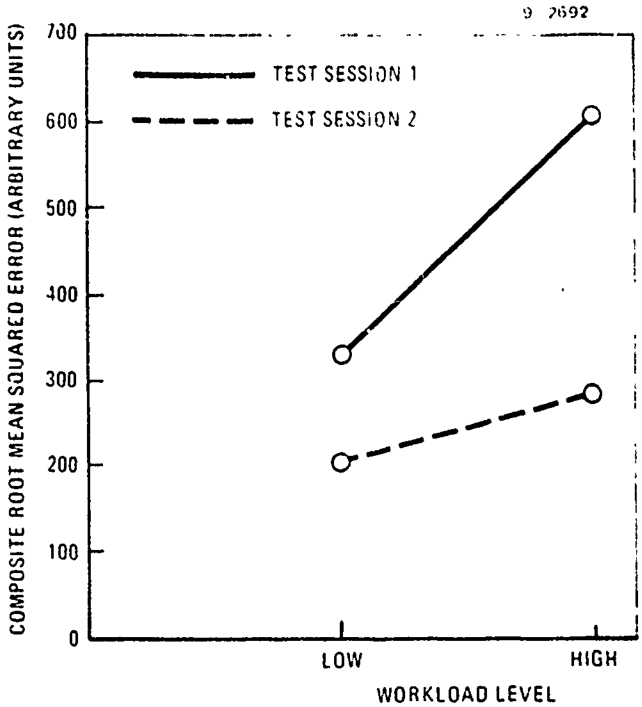

FIGURE 4–1 COMPOSITE MEASURES OF ROOT MEAN SQUARED ERROR (SUMMATIONS OF PITCH, BANK, AND AIRSPEED RMS ERRORS) AS A FUNCTION OF WORKLOAD LEVEL FOR THE FIRST AND SECOND TEST SESSIONS

to follow a commanded flight profile and maintain airspeed during the high workload trials, then there would be concomitant increases in the time needed to complete a concurrent target acquisition task. The term "resource" is a hypothetical construct (cf. Isreal et al., 1980a), used to explain performance variability in multiple-task situations. An individual has, at any point in time, a limited amount of resources which can be allocated to different, yet concurrent, tasks. If a primary task (in our case, following a commanded flight profile and maintaining a constant airspeed) places a particularly large demand on the available resources, then very few resources will remain for the performance of any other task (in our case, target acquisition).

Analysis of variance revealed that the main effect of the subjects variable was highly significant as well as extremely robust ($eta^2 > 38\%$); however, these data will not be discussed since we were not concerned with individual differences. The main effects of test session ($F = 17.53$, $df = 1/44$, $p < 0.001$) and workload level ($F = 6.30$, $df = 1/44$, $p < 0.001$) also were significant. As shown in

Figure 4-2, there was a substantial reduction in mean acquisition time from day 1 to day 2, due to continued practice on the task. Moreover, as we anticipated, acquisition times were shorter for the low workload than for the high workload condition. Planned comparisons indicated that acquisition times were shorter for the low workload than for the high workload condition on day 1 ($p < 0.01$), but did not differ significantly for the two workload levels on day 2. We conclude that on day 1, the demands imposed upon the pilots by the primary task were greater for the high workload than for the low workload condition. But by the second test session, the resources which had to be dedicated to the primary task during high workload and low workload trials were very similar.

4.3 Event-Related Potentials - Wickens (1980) has suggested that resources which must be devoted to perceptual encoding and information processing are functionally distinct from resources which must be devoted to the selection and execution of motor responses. The target acquisition data per se do not allow us to

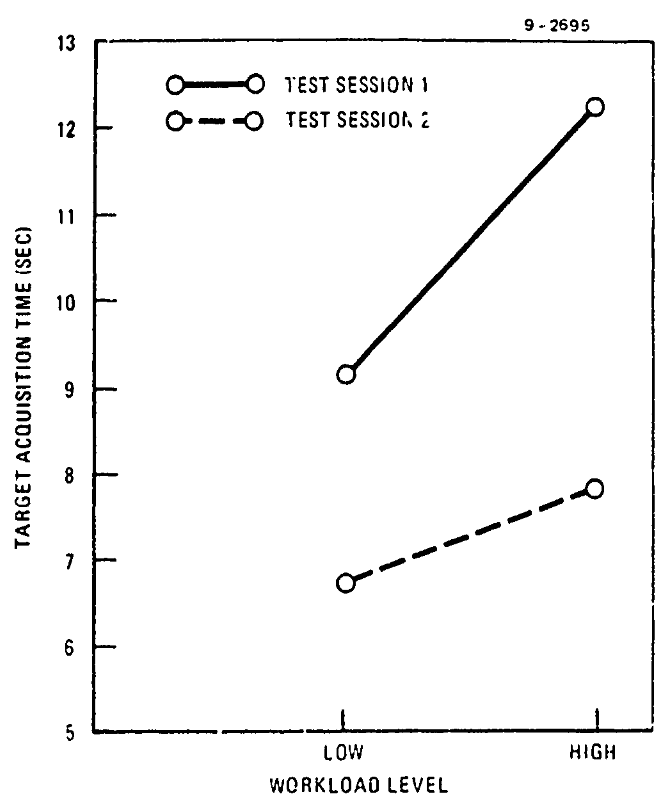

FIGURE 4–2 TARGET ACQUISITION TIME AS A FUNCTION OF WORKLOAD LEVEL FOR THE FIRST AND SECOND TEST SESSIONS

distinguish between demands made by the primary task on central processing resources and demands made by this task on response-related resources. The requirement to follow a commanded flight profile and maintain airspeed during high workload trials, expecially on day 1, could have imposed a heavy demand on available resources for coordinating and sequencing tracking responses, while only minimally drawing upon resources for perceptual encoding and information processing. Conversely, the primary task could have placed equal or even disproportionately high demands on central processing resources. Unfortunately, the longer target acquisition times we observed for high vs. low workload trials on day 1 would have resulted in either case, since behavioral reaction times are governed by central and response processes.

Donchin, McCarthy, Kutas, and Ritter (in press) have reviewed the literature which supports the position that the P_{300} component of the event-related potential is insensitive to manipulations of response selection or execution and. therefore, provides a specific index of the resources which must be allocated to encode and process information. As Donchin and his colleagues (Donchin, 1979, 1980; Isreal et al, 1980a, 1980b) have argued, the failure of background tonal events to elicit a large P_{300} in the context of a difficult primary task is due to an insufficient amount of central resources which remain in reserve for the processing of auditory information.

In the present investigation, artifact-free records of electrophysiological data were sorted for averaging according to subject, electrode placement, trial block within a session, test session, workload level, and threat avoidance stimulus (frequent vs. infrequent tone). Representative ensemble average ERPs computed for one pilot during the first and second test sessions are presented in Figure 4-3.

Two measures of the P_{300} component of ensemble average waveforms served as dependent variables in separate analyses. The first measure was P_{300} area, determined by summing the voltages of the post-stimulus time points between 300 and 420 msec, relative to the average pre-stimulus (100 msec) voltage level. As indicated in Figure 4-3, the measure of area between 300 and 420 msec does not encompass the entire P_{300}, particularly for infrequent tone presentations where positivity often extends well beyond 420 msec post-stimulus However, Donchin (Note 1, 1979) has

9 2694

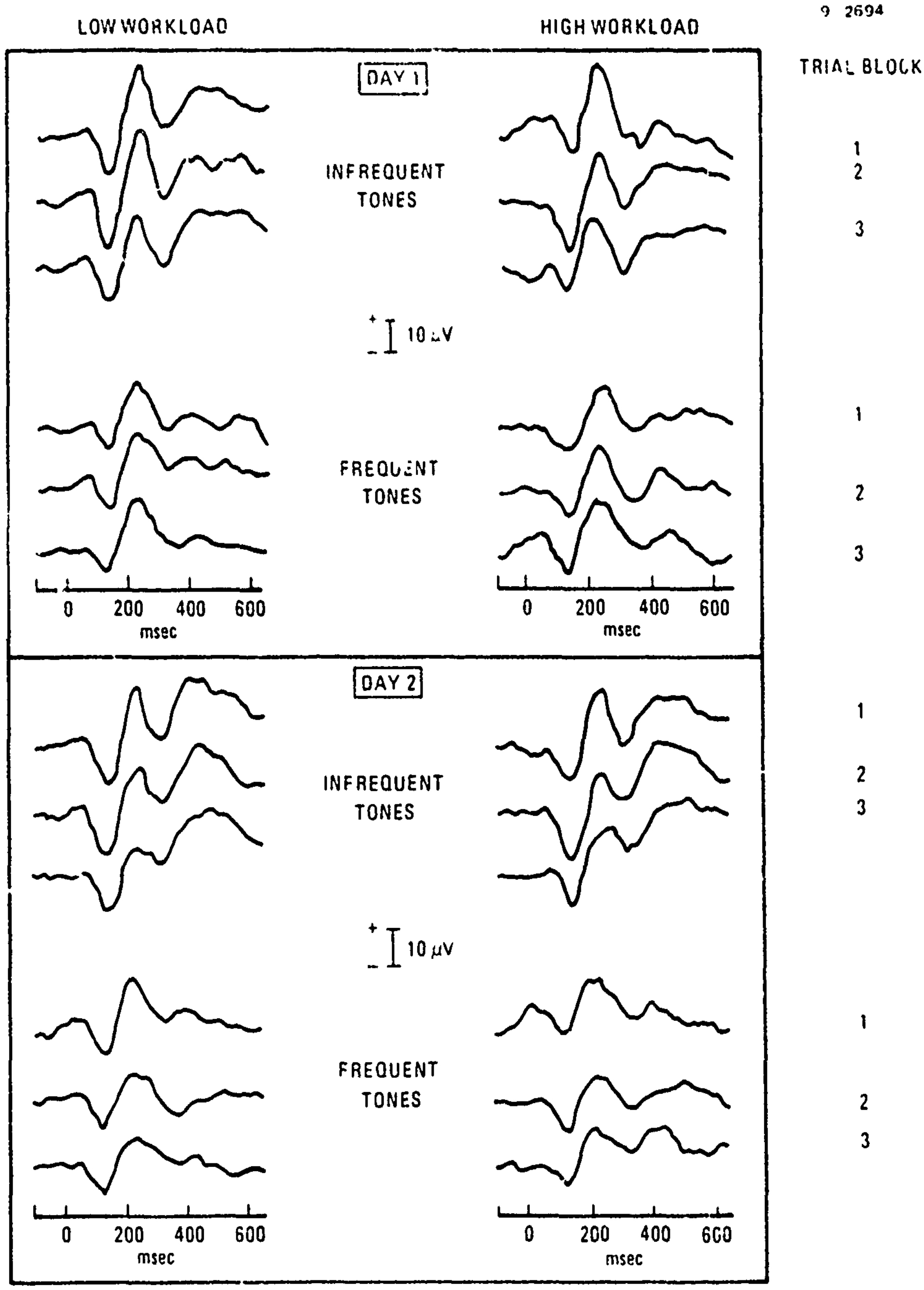

FIGURE 4–3 AVERAGE EVENT-RELATED POTENTIALS, ELICITED BY INFREQUENT AND FREQUENT TONES, ARE SHOWN FOR ONE SUBJECT AS A FUNCTION OF WORKLOAD LEVEL, TEST DAY, AND TRIAL BLOCK. BECAUSE OF THE SIMILARITIES BETWEEN LEFT PARIETAL, RIGHT PARIETAL AND MIDLINE PARIETAL WAVEFORMS, ONLY THE MIDLINE PARIETAL ERPs ARE PRESENTED

found that restricting the integration limits for area measurement provides a better metric of P_{300} for workload evaluation. The second measure was P_{300} latency, defined as that time point (within the limits of the total P_{300} component) associated with peak positivity.

For the measure of P_{300} area, analysis of variance indicated that the main effect of the subjects variable was again highly significant and robust (eta^2 >10%). Neither the main effect of electrode placement nor any interaction involving this variable were significant. Thus, the P_{300} components recorded from left and right parietal cortex, as well as from midline parietal cortex, were quite similar in magnitude. This was not too surprising. While many investigators (cf. Duncan-Johnson and Donchin, 1977; Squires, Donchin, Herning, and McCarthy, 1977) have reported that P_{300} is largest over parietal cortex during tasks requiring information processing, hemispheric differences are often not present (Donchin, Kutas, and McCarthy, 1977). The infrequent tone elicited a much larger P_{300} component than did the frequent tone ($F = 5.04$, $df = 1/328$, $p < 0.02$). As evident in Figure 4-3 a distinct P_{300} was often undiscernible in the waveform generated by frequent tonal events. We were most concerned with the influence of workload level and its interaction with threat avoidance stimulus and test session. There was a reliable main effect of workload level ($F = 4.61$, $df = 1/328$, $p < 0.03$), with high workload trials producing smaller area values than did low workload trials. Planned comparisons performed on the area measures for infrequent tonal events (pooling across electrode placement and trial block) revealed that P_{300} area was affected by experimental manipulations in the same manner as target acquisition time. Figure 4-4 shows that P_{300} area was substantially smaller for the high workload than for the low workload condition on day 1 ($p < 0.01$), but did not differ significantly for the two workload levels on day 2. In fact, P_{300} area for the high workload condition on day 1 was smaller than that for either workload condition on day 2 ($ps \leq 0.05$). We conclude that by the second test session, the central resources which had to be dedicated to the "high workload" primary task were reduced considerably and were similar to those demanded by the "low workload" primary task.

To obtain additional data which have a bearing on the validity of this conclusion, P_{300} latency measures were subjected to the same comparisons as were the area measures. Kutas, McCarthy, and Donchin (1977) and McCarthy and Donchin

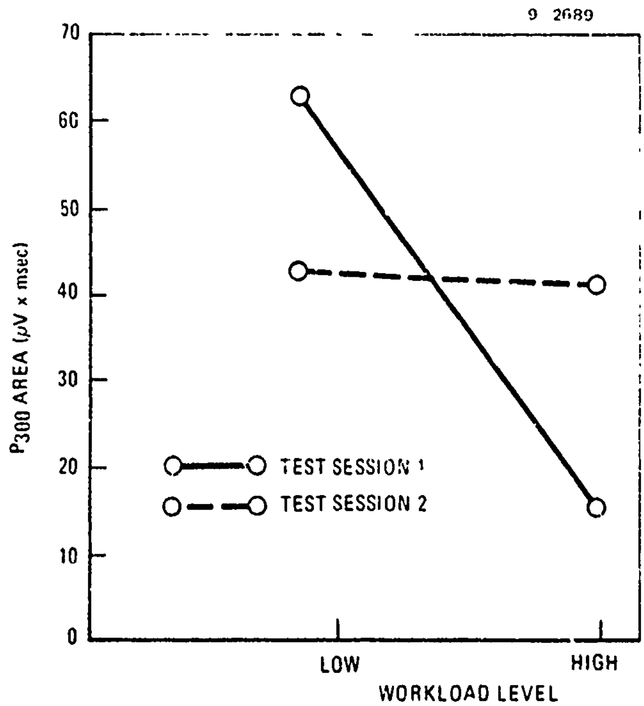

FIGURE 4–4 MAGNITUDE OF THE P_{300} COMPONENT AS A FUNCTION OF WORKLOAD LEVEL FOR THE FIRST AND SECOND TEST SESSIONS

(1981) have shown that P_{300} latency is sensitive to the duration of stimulus evaluation. Therefore, we predicted that increased primary task difficulty (and an increase in central resources allotted to the primary task) would be manifest in protracted P_{300} latencies to secondary tonal events. Analysis of variance indicated that the main effect of the subjects variable was extremely reliable and robust (eta^2 >43%). The infrequent tones produced longer latency P_{300}s than did the frequent tones ($F = 33.52$, $df = 1/328$, $p < 0.001$). There also was a marginally significant main effect of workload level ($F = 3.18$, $df = 1/328$, $p < 0.07$), with longer latencies associated with high workload than with low workload trials. Neither electrode placement nor trial block within a session influenced latency. By pooling across these variables and examining latencies for infrequent tonal events, P_{300} latency was found to be longer for the high workload than for the low workload condition on day 1 ($p < 0.05$), but did not differ for the two workload levels on day 2 (see Figure 4-5). Thus, we have shown that P_{300} area, P_{300} latency, and target acquisition times were affected by workload level and continued practice on concurrent flight tasks in a remarkably similar fashion.

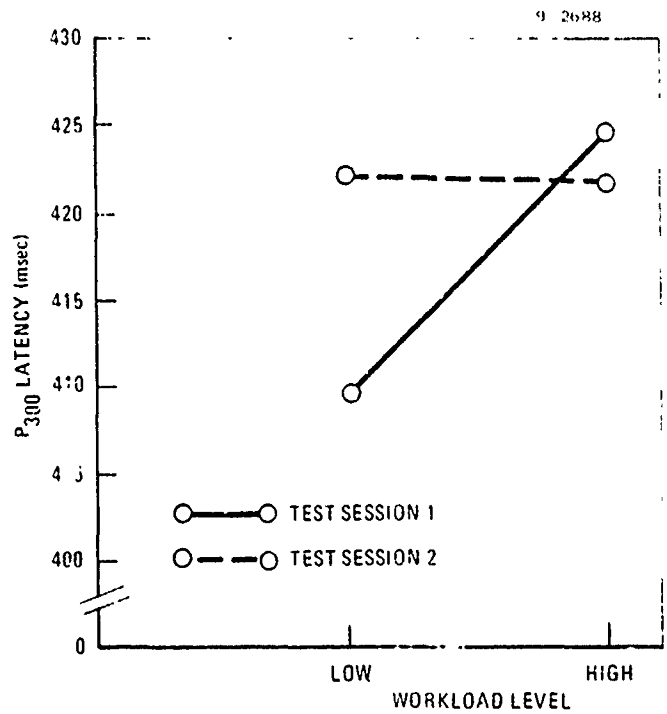

FIGURE 4–5 LATENCY OF THE P_{300} COMPONENT AS A FUNCTION OF WORKLOAD LEVEL FOR THE FIRST AND SECOND TEST SESSIONS

4.4 Electroencephalographic Activity - By sampling short periods (less than 750 msec) of electrocortical activity time-locked to a discrete event, we were able to isolate and then analyze changes in the transient waveform which were dependent upon the particular information processing requirements of the primary task. In the case of the ongoing EEG, however, the need to sample many seconds of data for estimating spectral features eliminates the possibility of establishing specific perceptual, cognitive, or motor correlates during complex task performance. Instead, measures of spectral intensity within limited frequency bands provide a direct indication of the tonic level of central nervous system (CNS) "activation" underlying the integration of perceptual, cognitive, and motor operations during multi-task performance. "Activation" is also a hypothetical construct that has been used to explain performance differences in complex task situations (cf. Gevins and Schaffer, 1980).

We assumed that as our pilots became more proficient in following a commanded flight profile while maintaining airspeed and in time-sharing between this primary task and the secondary target acquisition and threat avoidance tasks, there would

be concomitant changes in the spectral features of the EEG, denoting a reduction (with respect to initial level) in CNS activation. Moreover, we believed that our manipulation of workload, especially during the first test session, would influence the ease with which a pilot integrated primary and secondary task-related perceptual, cognitive, and motor functions. The more demanding the primary task and, consequently, the more difficult the integration process, then the greater the level of CNS activation we should find. Our assumption that the level of CNS activation would be reduced following continued experience with the primary task and with the time-sharing requirements is based upon a distinction made by Shiffrin and Schneider (1977) between two fundamental modes of processing: controlled and automatic. According to this view, controlled processing predominates as an individual begins to learn a complicated task or set of tasks. This mode of processing is thought to be highly demanding of resources which are available for perceptual encoding and information processing operations and which are available for the selection and execution of motor responses. Controlled processing may be described as conscious and quite effortful. It is usually slow, serial in nature, easily altered as problem solving strategies change, and is influenced greatly by task loading. For some aspects of a complicated task or set of tasks, an automatic mode of processing develops with continued training and overlearning. In contrast to controlled processing, automatic processing is relatively effortless, fast, parallel in nature, difficult to alter, and is much less affected by task loading. We assumed that the level of CNS activation would decrease as an automatic mode of processing began to govern some of the features of pilot performance.

Gevins and Schaffer (1980) have summarized a large number of studies reporting an inverse relationship between alpha-band power and CNS activation. They also reviewed some of their own work (Gevins, Zeitlin, Doyle, Yingling, Schaffer, Calloway, and Yeager, 1979) in which, relative to control conditions of minimal CNS activation, performance on spatial and analytic tasks was found to produce bilateral suppression of α-band power. The relationship between θ-band power and complex performance has been investigated recently in a series of experiments by Beatty and his colleagues (Beatty, Greenberg, Deibler, and O'Hanlon, 1974; Beatty and O'Hanlon, 1980; O'Hanlon and Beatty, 1977). They report that poor performance during vigilance tasks is associated with enhanced theta activity. As performance improves with continued practice, θ-band power is reduced.

We predicted that a reduction in the level of CNS activation and a greater role in performance for automatic processing would be associated with an increase in α-band power. We also expected that improvements in multi-task performance from Day 1 to Day 2 and across trial blocks within a session (from beginning to end) would produce decreases in θ-band power.

Mean spectral intensities were computed for each block of ten trials within a session and entered into the analyses of variance. Figure 4-6 illustrates the shift in the distribution of power from a rest period just before testing to actual performance conditions. Treatment effects on our measure of α-band power are now described. The main effect of the subjects variable was significant and

9 2693

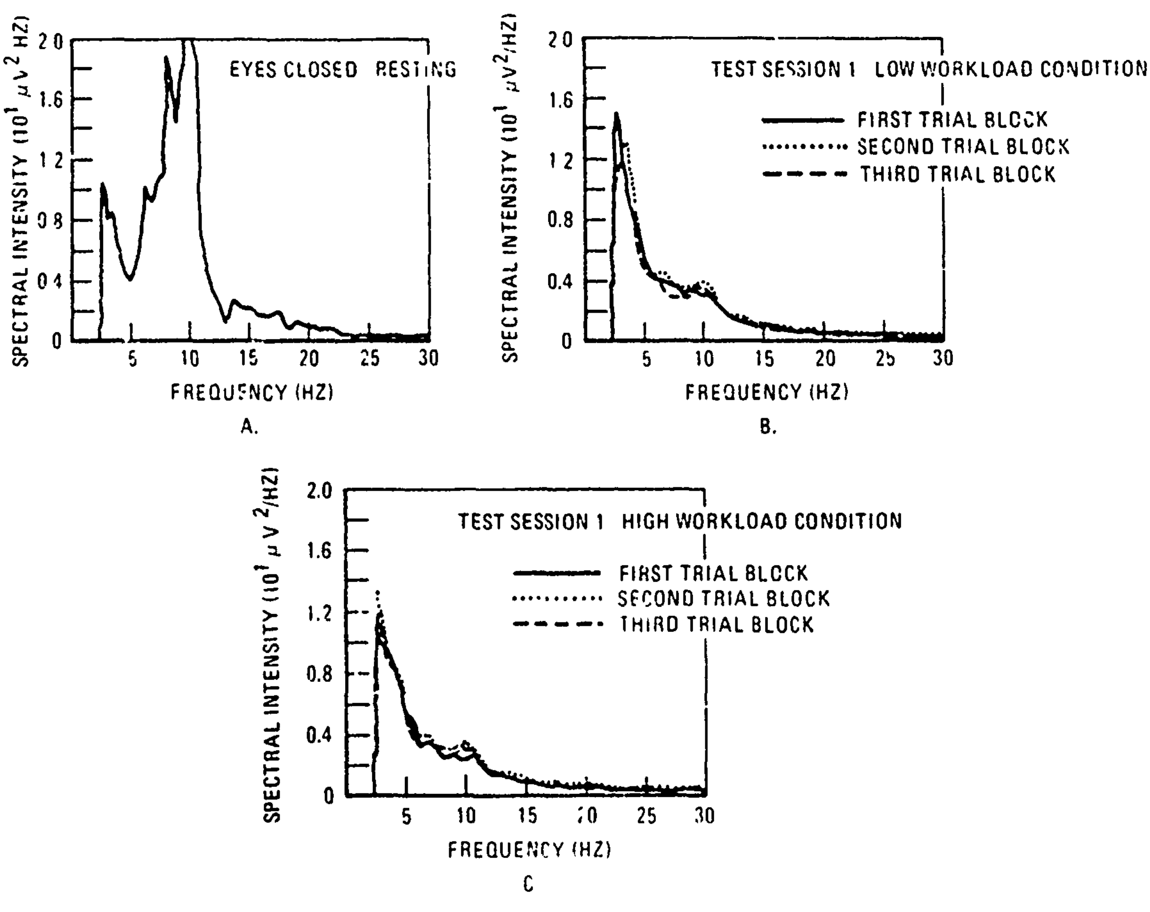

FIGURE 4–6 POWER SPECTRAL DENSITY PLOTS ARE SHOWN FOR ONE SUBJECT JUST PRIOR TO (PANEL A) AND DURING (PANELS B AND C) TEST SESSION ONE. ONLY THE MIDLINE PARIETAL DATA ARE PRESENTED.

very robust (eta^2 >60%). Further, the main effects of workload level ($F = 11.40$, $df = 1/144$, $p < 0.001$), electrode placement ($F = 4?.91$, $df = 2/144$, $p < 0.001$), and trial block within a session ($F = 7.62$, $df = 2/144$, $p < 0.001$) also were significant. As we anticipated, high workload trials were associated with a reduction in α-band power, indicating greater CNS activation for this condition than for the low workload condition. Planned comparisons revealed that this distinction was reliable during both test sessions (ps <0.05), in contrast to the behavioral and ERP data for which workload level ceased to be a significant factor by the second session. The significant main effect of electrode placement was due to greater α-band power recorded from midline parietal cortex than from either left or right parietal cortex (ps <0.05). No difference in power was found between the left and right parietal cortex, thus supporting the earlier conclusion of Gevins et al. (1979) that CNS activation during complex task performance produces a bilateral suppression of α-band power.

Unexpectedly, the main effect of the test session variable was unreliable. That is, there was no appreciable reduction in our measure of CNS activation from the first to the second session, even though primary and secondary task performance did improve dramatically by the second test session. However, Figure 4-7 shows the changes in α-band power that occurred as a function of trial block within both test sessions. (We have pooled across workload level because of the similarities in the plots when these data are graphed as in Figure 4-7.) Significant increases in α-band power were found from the first to the third block of ten trials on both test days (ps <0.05). We should note that although the main effect of trial block did not achieve statistical significance for either the RMS error data or the target acquisition latencies, nonetheless, there were consistent trends of increased accuracy and reduced latency from the first to the third trial block. Therefore, we conclude that this apparent reduction in CNS activation is consistent with the hypothesis that continued practice on the simulated mission led to the development of an automatic mode of processing for certain aspects of the flight tasks.

Turning to the θ-band spectral data, the main effect of the subjects variable again was significant and highly robust (eta^2 >40%). The analysis of variance indicated that the main effects of workload level ($F = 9.38$, $df = 1/144$, $p < 0.002$) and electrode placement ($F = 55.73$, $df = 2/144$, $p < 0.001$) also were significant.

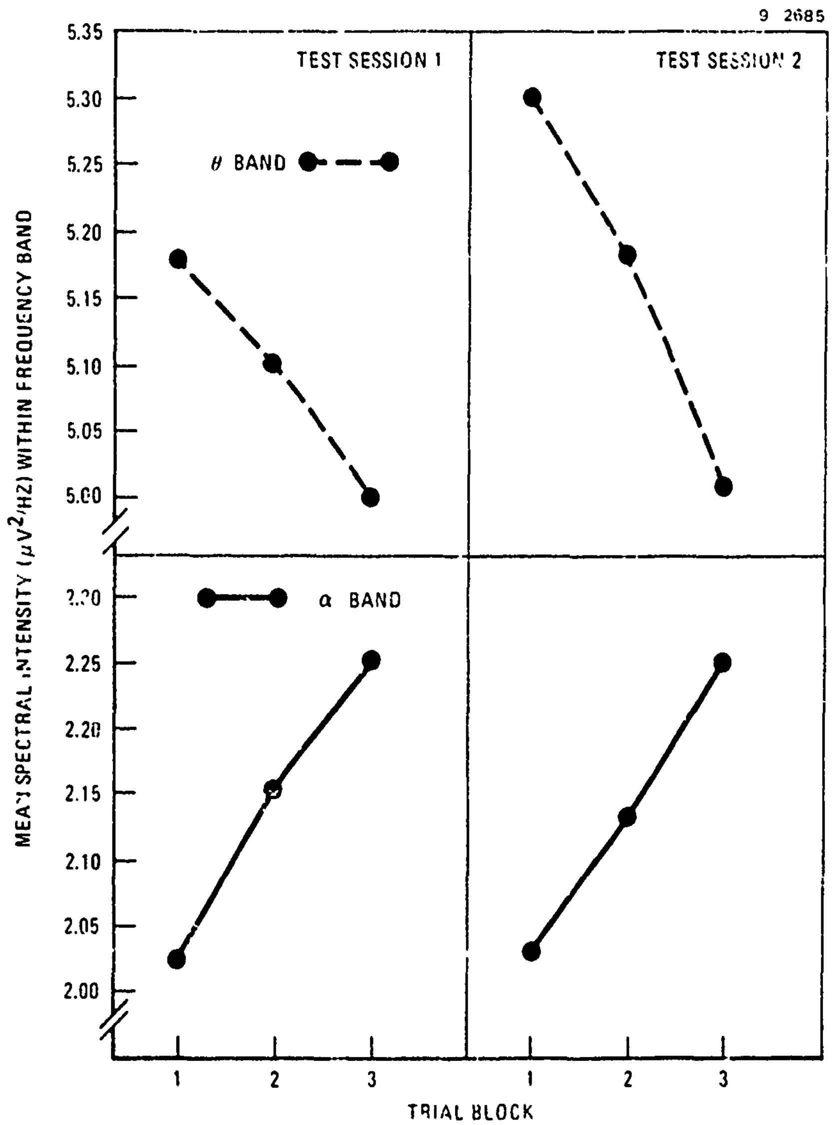

FIGURE 4–7 MEAN α – AND θ -BAND POWER AS A FUNCTION OF TRIAL BLOCK FOR THE FIRST AND SECOND TEST SESSIONS

As was the case for α-band power, greater θ-band power was recorded over midline parietal cortex than over either left or right parietal cortex ($ps < 0.05$). In addition, the level of θ-band power was higher over right than over left parietal cortex ($p < 0.05$). To the extent that poorer performance in following a commanded flight profile and maintaining airspeed was found during the high workload trials on the first test session, we expected to see greater θ-band power on these trials as well. Inexplicably, reduced θ-band power was associated with high workload

trials. Moreover, since primary and secondary task performance improved reliably from the first to the second test session, we assumed that θ-band power would decline by the second session. However, as we reported for the α-band data, the main effect of test session was not significant.

Figure 4-7 presents the changes in θ-band power that were found as a function of trial block within both test sessions. As expected, the reliable decreases in θ-band power from the first to the third block of trials on both test days ($ps < 0.05$) were exactly opposite to the significant increases observed for α-band power. Taken together, we conclude that these data are consistent with the hypothesis that an effortful, conscious control of processing was replaced during the latter trials within each session by an automatic mode of processing for at least some features of the multiple flight tasks.

5.0 RECOMMENDATIONS

The results of this investigation are encouraging. We have taken a physiological method used to assess mental workload in laboratory settings and applied it successfully to a situation in which individuals performed operationally relevant tasks. We recommend that this method be studied in the context of even more demanding simulations. In this way, appropriate data will be made available to specify its relative strengths and weaknesses on a more practical basis.

6.0 REFERENCES

Reference Notes

1. Donchin, E. Personal communication, 1979.

References

Beatty, J., Greenberg, A., Deibler, W. P., and O'Hanlon, J. F. Operant control of occipital theta rhythm affects performance on a radar monitoring task. Science, 1974, 183, 871-873.

Beatty, J., and O'Hanlon, J. F. Operant control of posterior theta rhythms and vigilance performance: Repeated treatments and transfer of training. In N. Birbaumer and K. D. Kimmel (Eds.), Biofeedback and Self-Regulation. Hillsdale, New Jersey: Erlbaum Press, 1980.

Donchin, E. Event-related potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked Brain Potentials and Behavior. New York: Plenum Press, 1979.

Donchin, E. Event-related potentials: Inferring cognitive activity in operational settings. In F. E. Gomer (Ed.), Biocybernetic Applications for Military Systems: Proceedings of the DARPA Conference, Chicago, 1978. McDonnell Douglas Astronautics Company - St. Louis Division Technical Report MDC E2191, 1980.

Donchin, E., Kutas, M., and McCarthy, G. Electrocortical indices of hemispheric utilization, In S. Harnad, R. Doty, L. Goldstein, J. Jaynes, and G. Krauthamer (Eds.), Lateralization in the Nervous System. New York: Academic Press, 1977.

Donchin, E., McCarthy, G., Kutas, M., and Ritter, W. Event-related brain potentials in the study of consciousness. In J. M. Davidson, G. E. Schwartz, and D. Shapiro (Eds.), Consciousness and Self Regulation Vol. III. New York: Plenum Press, (in press).

Duncan-Johnson, C. C., and Donchin, E. On quantifying surprise: The variation in event-related potentials with subjective probability. Psychophysiology, 1977, 14, 456-467.

Gevins, A. S., and Schaffer, R. E. A critical review of electroencephalographic (EEG) correlates of higher cortical functions. CRT Critical Reviews in Bioengineering, 1980, 4, 113-164.

Gevins, A., Zeitlin, G., Doyle, J., Yingling, C., Schaffer, R., Calloway, E., and Yeager, C. Electroencephalogram correlates of higher cortical functions. Science, 1979, 203, 665-667.

Gomer, F. E. Physiological monitoring and the concept of adaptive systems. In J. Moraal and K. F. Kraiss (Eds.), Manned Systems Design: Methods, Equipment, and Applications. New York: Plenum Press, (in press).

Gomer, F. E., Beideman, L. R., and Levine, S. H. The application of biocybernetic techniques to enhance pilot performance during tactical missions. McDonnell Douglas Astronautics Company-St. Louis Division Technical Report MDC E2046, 1979.

Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E. P_{300} and tracking difficulty: Evidence for multiple resources in dual-task performance. Psychophysiology, 1980a, 17, 259-273.

Isreal, J. B., Wickens, C. D., Chesney, G. L., and Donchin, E. The event-related brain potential as an index of display-monitoring workload. Human Factors, 1980b, 22, 211-224.

Johannsen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., and Wickens, C. Final report of the experimental psychology groups. In N. Moray (Ed.), Mental Workload: Its Theory and Measurement. New York: Plenum Press, 1979.

Knoop, P. A., and Welde, W. L. Automated pilot performance assessment in the T-37. Air Force Human Resources Laboratory Technical Report TR-72-6, 1973.

Kutas, M., McCarthy, G., and Donchin, E. Augmenting mental chronometry: The P_{300} as a measure of stimulus evaluation time. Science, 1977, 197, 792-795.

McCarthy, G., and Donchin, E. A metric for thought: A comparison of P_{300} latency and reaction time. Science, 1981, 211, 77-79.

Ogden, G. D., Levine, J. M., and Eisner, E. J. Measurement of workload by secondary tasks. Human Factors, 1979, 21, 529-548.

O'Hanlon, J. F., and Beatty, J. Concurrence of electroencephalographic and performance changes during a simulated radar watch and some implications for the arousal theory of vigilance. In R. Mackie (Ed.), Vigilance: Theory, Operational Performance, and Physiological Correlates. New York: Plenum Press, 1977.

Shiffrin, R. M., and Schneider, W. Controlled and automatic human information processing II: Perceptual learning, automatic attending, and a general theory. Psychological Review, 1977, 84, 127-190.

Spyker, D. A., Stackhouse, S. P., Khalafalla, A. S., and McLane, R. C. Development of techniques for measuring pilot workload. National Aeronautics and Space Administration Contractor Report CR 1888, 1971.

Squires, K. C., Donchin, E., Herning, R. I., and McCarthy, G. On the influence of task relevance and stimulus probability on event-related potential components. Electroencephalography and Clinical Neurophysiology, 1977, 42, 1-14.

Thatcher, R. W., and John, E. R. Functional Neuroscience Volume I: Foundations of Cognitive Processes. New York: John Wiley and Sons, 1977.

Warner, H. D., Drenner, T. G., and Curtin, J. G. Manual control in target tracking tasks as a function of control characteristics. McDonnell Douglas Astronautics Company-St. Louis Division Technical Report MDC E1433, 1976.

Wickens, C. D. The structure of attentional resources. In R. Nickerson (Ed.), Attention and Performance VIII. Englewood Cliffs, NJ: Lawrence Erlbaum Assoc., 1980.

Williges, R. C., and Wierwille, W. W. Behavioral measures of aircrew mental workload. Human Factors, 1979, 21, 549-574.